united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

ICTP 40th Anniversary

SMR 1595 - 5

Joint DEMOCRITOS - ICTP School on
CONTINUUM QUANTUM MONTE CARLO METHODS
12 - 23 January 2004

VARIATIONAL MONTE CARLO FOR ATOMS AND MOLECULES

Claudia FILIPPI
Instituut-Lorentz, Universiteit Leiden, The Netherlands

Variational Monte Carlo for atoms and molecules

Claudia Filippi

Instituut-Lorentz, Universiteit Leiden, The Netherlands

1. Metropolis algorithm
    - Choice of proposal matrix
2. Trial wave function
    - Spin projection
    - Cusp conditions
    - Jastrow factor
    - Static correlation
3. Optimization of wave function
    - Variance minimization
    - Energy minimization
4. Correlated sampling
    - Computation of potential energy difference

Electronic structure calculations

First-principle description

Molecules, solids → Collection of ions + electrons

$$\downarrow$$

Only input: Z_α, N_α

Work in the Born-Oppenheimer approximation

→ Separate nuclear and electronic degrees of freedom

Solve Schrödinger equation for electrons in ionic field

$$\mathcal{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_i v_{\text{ext}}(\mathbf{r}_i) + \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}$$

What do we want to compute?

Fermionic ground state and low-lying excited states

Evaluate expectation values $\quad \dfrac{\langle\Psi_n|\mathcal{O}|\Psi_n\rangle}{\langle\Psi_n|\Psi_n\rangle}$

Electronic structure: possible approaches

(a) Density functional theory methods

Finite and extended systems

Approximate treatment of exchange-correlation

(b) Quantum chemistry methods

Post Hartree-Fock wave function methods

CI MCSCF CC ...

Accurate on small systems

(c) Quantum Monte Carlo techniques

Fully-correlated calculations

Stochastic solution of the Schrödinger equation

Most accurate benchmarks for medium-large systems: 1^{st}-2^{nd}-row clusters with N_{atom}=20–50 and solids, where QC methods are difficult to apply

Quantum Monte Carlo

- Variational Monte Carlo

Monte Carlo as a way of evaluating integrals

Consider many-body wave function $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$

Compute expectation value of $\mathcal{O}$ operator ($\mathcal{H}$, n ...)

$$\langle \mathcal{O} \rangle_{\mathrm{VMC}} = \frac{\langle \Psi | \mathcal{O} | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$

Why should we use Monte Carlo integration?

⇒ Freedom in functional form of Ψ

- Projection Monte Carlo Methods

– Diffusion Monte Carlo

(Grimm & Storer, Anderson, Ceperley, 1971-1980)

– Domain Green Function Monte Carlo

(Kalos, 1974)

– Other variants, e.g. Reptation MC

(Baroni, Moroni, 1998)

Expectation values in Monte Carlo methods

Probability distribution ρ (continuous or discrete)

Monte Carlo to compute expectation values as

$$\frac{\int \mathrm{d}\mathbf{R}\, O(\mathbf{R})\rho(\mathbf{R})}{\int \mathrm{d}\mathbf{R}\, \rho(\mathbf{R})} \approx \frac{1}{M}\sum_{i=1}^{M} O(\mathbf{R}_i)$$

Configurations $\mathbf{R}_i$ are distributed as $\rho(\mathbf{R})/\int \mathrm{d}\mathbf{R}\, \rho(\mathbf{R})$

In variational Monte Carlo

$$\langle \mathcal{O} \rangle_{\mathrm{VMC}} = \frac{\langle \Psi | \mathcal{O} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \int \mathrm{d}\mathbf{R}^{3\mathrm{N}} \left(\frac{\mathcal{O}\Psi}{\Psi}\right)_{\mathbf{R}} \boxed{\frac{|\Psi(\mathbf{R})|^2}{\int \mathrm{d}\mathbf{R}^{3\mathrm{N}} |\Psi(\mathbf{R})|^2}}$$

$$\mathbf{R} = (\mathbf{r}_1, \ldots, \mathbf{r}_N), \quad \rho(\mathbf{R}) = |\Psi(\mathbf{R})|^2, \quad O(\mathbf{R}) = \left(\frac{\mathcal{O}\Psi}{\Psi}\right)_{\mathbf{R}}$$

$$\langle \mathcal{O} \rangle_{\mathrm{VMC}} \approx \frac{1}{M}\sum_{i=1}^{M} \left(\frac{\mathcal{O}\Psi}{\Psi}\right)_{\mathbf{R}_i}$$

We need a means to sample ρ

Metropolis Algorithm

Aim: Obtain a set of $\{\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_M\}$ distributed as a given $\rho(\mathbf{R})$

Let us generate a Markov chain:

- Start from arbitrary initial state $\mathbf{R}_i$
- Use stochastic transition matrix $M(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})$

$$M(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) \geq 0 \qquad \sum_{\mathbf{R}_\mathsf{f}} M(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = 1.$$

  as probability of making transition $\mathbf{R}_\mathsf{i} \to \mathbf{R}_\mathsf{f}$

- Evolve the system by repeated application of M

To sample ρ

M must satisfy stationarity condition:

$$\sum_i M(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\, \rho(\mathbf{R}_\mathsf{i}) = \rho(\mathbf{R}_\mathsf{f}) = \sum_i M(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})\, \rho(\mathbf{R}_\mathsf{f}) \quad \forall\, \mathbf{R}_\mathsf{f}$$

$\Rightarrow$ If we start with ρ, we continue to sample ρ

Stationarity + stochastic property of M + ergodicity

$\Rightarrow$ Any initial distribution evolves to ρ

$\boxed{\text{How do we construct } M \text{ in practice?}}$

M must satisfy stationarity condition:

$$\sum_i M(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\ \rho(\mathbf{R}_\mathsf{i}) = \sum_i M(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})\ \rho(\mathbf{R}_\mathsf{f}) \quad \forall\ \mathbf{R}_\mathsf{f}$$

- Impose $\boxed{\text{detailed balance}}$ condition

$$\boxed{M(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\ \rho(\mathbf{R}_\mathsf{i}) = M(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})\ \rho(\mathbf{R}_\mathsf{f})}$$

  $\underline{\text{Sufficient}}$ but $\underline{\text{not necessary}}$ condition

- Write M as $\boxed{\text{proposal } T \times \text{acceptance } A}$

$$M(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\ T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})$$

  M and T are stochastic matrices but A is not

  Detailed balance is now:

$$A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\ T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\ \rho(\mathbf{R}_\mathsf{i}) = A(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})\ T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})\ \rho(\mathbf{R}_\mathsf{f})$$

  or

$$\boxed{\frac{A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}{A(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})} = \frac{T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})\ \rho(\mathbf{R}_\mathsf{f})}{T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\ \rho(\mathbf{R}_\mathsf{i})}}$$

Choice of acceptance matrix A

For a given choice of T, infinite choices of A satisfy

$$\frac{A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}{A(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})} = \frac{T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})\ \rho(\mathbf{R}_\mathsf{f})}{T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\ \rho(\mathbf{R}_\mathsf{i})}$$

Any function $A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = F\left(\frac{T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})\ \rho(\mathbf{R}_\mathsf{f})}{T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\ \rho(\mathbf{R}_\mathsf{i})}\right)$

with $F(x)/F(1/x) = x$ will do

Choice by Metropolis *et al.* maximizes the acceptance

$$\boxed{A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = \min\left\{1, \frac{T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})\ \rho(\mathbf{R}_\mathsf{f})}{T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})\ \rho(\mathbf{R}_\mathsf{i})}\right\}}$$

Note: $\rho(\mathbf{R})$ does not have to be normalized

Original Metropolis method

Symmetric proposal matrix $T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f}) = T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})$

$$A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = \min\left\{1, \frac{\rho(\mathbf{R}_\mathsf{f})}{\rho(\mathbf{R}_\mathsf{i})}\right\}$$

Aim: Obtain a set of $\{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_M\}$ distributed as a given $\rho(\mathbf{R})$

Operationally

1. Pick a starting $\mathbf{R}$ and evaluate $\rho(\mathbf{R})$
2. Choose $\mathbf{R}'$ at random
3. If $\rho(\mathbf{R}') \geq \rho(\mathbf{R})$, move accepted
   → put $\mathbf{R}'$ in the set
4. If $\rho(\mathbf{R}') < \rho(\mathbf{R})$, move accepted with $p = \dfrac{\rho(\mathbf{R}')}{\rho(\mathbf{R})}$

   To do this, pick a random number $\chi \in [0, 1]$:

   a) If $\chi < p$, move accepted
      → put $\mathbf{R}'$ in the set

   b) If $\chi > p$, move rejected
      → put another entry of $\mathbf{R}$ in the set

Choice of proposal matrix T (1)

Metropolis method → Points sequentially correlated

Aim → Achieve fastest evolution of the system

⇒ High acceptance + large proposed moves

⇒ Find optimal T with high acceptance + large moves

Original Metropolis method $T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f}) = T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})$

In general

$$A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) = \min\left\{1, \frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ \rho(\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ \rho(\mathbf{R}_\mathrm{i})}\right\}$$

Use <u>freedom</u> in the choice of T to make

$$\frac{T(\mathbf{R}_\mathrm{i}|\mathbf{R}_\mathrm{f})\ \rho(\mathbf{R}_\mathrm{f})}{T(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i})\ \rho(\mathbf{R}_\mathrm{i})} \approx 1 \quad \Rightarrow \quad A(\mathbf{R}_\mathrm{f}|\mathbf{R}_\mathrm{i}) \approx 1$$

and <u>reduce autocorrelation time</u> of desired observable

<u>Note</u>: we need to be able to sample T directly

Choice of proposal matrix T (2)

C. Umrigar, Phys. Rev. Lett. **71**, 408 (1993)

Rewrite proposal matrix T as

$$T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = \frac{S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}{\int \mathrm{d}\mathbf{R}_\mathsf{f}\, S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})} = \frac{S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}{I(\mathbf{R}_\mathsf{i})}$$

with $I(\mathbf{R}_\mathsf{i}) = \int \mathrm{d}\mathbf{R}_\mathsf{f}\, S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) \;\Rightarrow\; \int \mathrm{d}\mathbf{R}_\mathsf{f}\, T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = 1$

$$\Rightarrow \quad \frac{T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})}{T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}\frac{\rho(\mathbf{R}_\mathsf{f})}{\rho(\mathbf{R}_\mathsf{i})} = \frac{I(\mathbf{R}_\mathsf{i})}{I(\mathbf{R}_\mathsf{f})}\frac{S(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})}{S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}\frac{\rho(\mathbf{R}_\mathsf{f})}{\rho(\mathbf{R}_\mathsf{i})}$$

- If $\boxed{I(\mathbf{R}_\mathsf{i}) = I(\mathbf{R}_\mathsf{f})}$ for all $\mathbf{R}_\mathsf{f}$ accessible from $\mathbf{R}_\mathsf{i}$

$$S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) \sim \rho(\mathbf{R}_\mathsf{f}) \quad \Rightarrow \quad \frac{T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})}{T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}\frac{\rho(\mathbf{R}_\mathsf{f})}{\rho(\mathbf{R}_\mathsf{i})} \approx 1$$

$$\Leftrightarrow \quad \frac{A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}{A(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})} \approx 1$$

Usually, not possible to find approximation to ρ over all domain of ρ which can be sampled directly

$\boxed{\text{Choice of proposal matrix } T}$ (3)

- Usually, we choose $\boxed{S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) \neq 0 \text{ for } \mathbf{R}_\mathsf{f} \in D(\mathbf{R}_\mathsf{i})}$ with $D(\mathbf{R}_\mathsf{i})$ a domain of volume $\Omega(\mathbf{R}_\mathsf{i})$ around $\mathbf{R}_\mathsf{i}$ $\Rightarrow$ Proposed moves are in domain $D(\mathbf{R}_\mathsf{i})$

Now, $I(\mathbf{R}_\mathsf{f}) \neq I(\mathbf{R}_\mathsf{i})$ and

$$I(\mathbf{R}_\mathsf{i}) = \int \mathrm{d}\mathbf{R}_\mathsf{f}\, S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) \approx S(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{i})\Omega(\mathbf{R}_\mathsf{i}) \Rightarrow$$

$$\frac{T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})}{T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}\frac{\rho(\mathbf{R}_\mathsf{f})}{\rho(\mathbf{R}_\mathsf{i})} \approx \frac{\Omega(\mathbf{R}_\mathsf{i})}{\Omega(\mathbf{R}_\mathsf{f})}\frac{S(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{i})}{S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{f})}\frac{S(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})}{S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}\frac{\rho(\mathbf{R}_\mathsf{f})}{\rho(\mathbf{R}_\mathsf{i})}$$

Choosing $S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = g(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) \Big/ \sqrt{\Omega(\mathbf{R}_\mathsf{f})}$ and

$$g(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) \sim \sqrt{\rho(\mathbf{R}_\mathsf{f})} \quad \Rightarrow \quad \frac{T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})}{T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}\frac{\rho(\mathbf{R}_\mathsf{f})}{\rho(\mathbf{R}_\mathsf{i})} \approx 1$$

$$\Leftrightarrow \quad \frac{A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}{A(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})} \approx 1$$

Choice of proposal matrix T

- If Δ is the linear dimension of domain $D(\mathbf{R}_\mathsf{i})$

$$\frac{A(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})}{A(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})} = \frac{T(\mathbf{R}_\mathsf{i}|\mathbf{R}_\mathsf{f})}{T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})} \frac{\rho(\mathbf{R}_\mathsf{f})}{\rho(\mathbf{R}_\mathsf{i})} \approx 1 - \mathcal{O}(\Delta^m)$$

$m = 1$ $\quad S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})$ symmetric

$m = 1$ $\quad S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) \sim \rho(\mathbf{R}_\mathsf{f})$

$m = 2, 3$ $\quad \nabla \ln g(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = \nabla \ln \sqrt{\rho(\mathbf{R}_\mathsf{f})}$ at $\mathbf{R}_\mathsf{f} = \mathbf{R}_\mathsf{i}$

Metropolis algorithm in electronic structure theory

Calculate quantum mechanical expectation values

For example, the total energy is given by

$$\begin{aligned}
\langle \mathcal{H} \rangle_{\mathrm{VMC}} &= \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \\
&= \int \mathrm{d}\mathbf{R}^{3\mathrm{N}} \left(\frac{\mathcal{H}\Psi}{\Psi} \right)_{\mathbf{R}} \boxed{\frac{|\Psi(\mathbf{R})|^2}{\int \mathrm{d}\mathbf{R}^{3\mathrm{N}} |\Psi(\mathbf{R})|^2}} \\
&= \frac{1}{M} \sum_{i=1}^{M} \left(\frac{\mathcal{H}\Psi}{\Psi} \right)_{\mathbf{R}_i} \\
&= \frac{1}{M} \sum_{i=1}^{M} E_{\mathrm{L}}(\mathbf{R}_i)
\end{aligned}$$

Note: If $\Psi \rightarrow$ eigenfunction, $E_L(\mathbf{R})$ does not fluctuate

$\Rightarrow$ Importance of optimizing trial wave function

⋆ $\boxed{\text{This afternoon}}$

1. <u>Simple Metropolis</u> $(m=1)$

   $S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})$ is constant in a box centered in $\mathbf{R}_\mathsf{i}$

   In each of the $3N$ dimensions, sample uniformely

   $$d\,x = x_\mathsf{f} - x_\mathsf{i} \in \left[-\frac{\Delta}{2}, \frac{\Delta}{2}\right]$$

2. <u>Directed Metropolis</u> $(m=2)$

   $S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i})$ is a linear approximation to $\Psi(\mathbf{R}_\mathsf{f})$ at $\mathbf{R}_\mathsf{i}$

   $$S(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = \prod_{k=1}^{3N} \left\{1 + (x_{k,\mathsf{f}} - x_{k,\mathsf{i}}) \times \min\left[|\mathbf{V}_k(\mathbf{R}_\mathsf{i})|, \frac{2}{\Delta}\right] \times \mathrm{sign}[\mathbf{V}_k(\mathbf{R}_\mathsf{i})]\right\}$$

   with $x_{k,\mathsf{f}} - x_{k,\mathsf{i}} \in \left[-\frac{\Delta}{2}, \frac{\Delta}{2}\right]$ and $\mathrm{V}(\mathbf{R}_\mathsf{i}) = \frac{\nabla\Psi(\mathbf{R}_\mathsf{i})}{\Psi(\mathbf{R}_\mathsf{i})}$

3. <u>Motivated by diffusion Monte Carlo</u> $(m=2)$

   $$T(\mathbf{R}_\mathsf{f}|\mathbf{R}_\mathsf{i}) = \frac{1}{(2\pi\tau)^{3N/2}} \exp\left[-\frac{(\mathbf{R}_\mathsf{f} - \mathbf{R}_\mathsf{i} - \bar{\mathbf{V}}(\mathbf{R}_\mathsf{i})\tau)^2}{2\tau}\right]$$

   Limit V as $\bar{\mathbf{V}} = \frac{\sqrt{1+2\,a\,V^2\,\tau} - 1}{a\,V^2\,\tau}\mathbf{V}$

Autocorrelation time

Run of N Monte Carlo steps = N_b blocks $\times$ N_s steps

We have N measurements of E_L

$\bar{E}$ = average of E_L

σ = rms fluctuations of individual E_L

σ_b = rms fluctuations of block averages of E_L

Effectively, N/T_corr independent measurements of E_L

Define T_corr as

$$\mathrm{err}(\bar{E}) = \frac{\sigma}{\sqrt{N_b \times N_s}}\sqrt{T_\mathrm{corr}} = \frac{\sigma_b}{\sqrt{N_b}}$$

$\Rightarrow$ $\boxed{T_\mathrm{corr} = N_s \left(\frac{\sigma_b}{\sigma}\right)^2}$ where we chose $N_s \gg T_\mathrm{corr}$

Autocorrelation time and acceptance versus step size

Example: Be, 4 determinants + simple Jastrow factor

$$E_{\text{VMC}} = -14.9581(3)\ \text{H},\ \sigma_{\text{VMC}} = 0.35\ \text{H}$$

1. Simple Metropolis

Δ	T_{corr}	$\bar{A}$
1.00	41	0.17
0.75	21	0.28
0.50	17	0.46
0.20	45	0.75

2. Directed Metropolis

Δ	T_{corr}	$\bar{A}$
1.20	21	0.38
1.00	11	0.52
0.75	6	0.72
0.50	8	0.88
0.20	34	0.99

3. Drift-diffusion transition

τ	T_{corr}	$\bar{A}$
0.100	13	0.42
0.050	7	0.66
0.020	8	0.87
0.010	14	0.94

Shortcomings of Metropolis algorithms 1-2-3

- No distinction between core and valance electrons
  $\Rightarrow$ Core electrons set the length scales
- Use of cartesian coordinates
  $\Rightarrow$ Derivative discontinuity of Ψ at nuclei
- All-electron versus single-electron move

Better algorithms can achieve $T_{\text{corr}} = 1 - 2$

Trial wave function

Traditional quantum chemistry

$$\Psi(\mathbf{x}_1,\dots,\mathbf{x}_N)=\Psi(\mathbf{r}_1,\sigma_1,\dots,\mathbf{r}_N,\sigma_N) \quad \text{with} \quad \sigma_i=\pm 1$$

Hartree-Fock

$$\begin{tikzcd}
\Psi(\mathbf{x}_1,\dots,\mathbf{x}_N) \arrow[rr] \arrow[dr] & & D_{\mathrm{HF}}=\begin{vmatrix} \psi_1(\mathbf{x}_1) & \dots & \psi_1(\mathbf{x}_N) \\ \vdots & & \vdots \\ \psi_N(\mathbf{x}_1) & \dots & \psi_N(\mathbf{x}_N) \end{vmatrix} \arrow[dl] \\
& c_0 D_{\mathrm{HF}}+c_1 D_1+c_2 D_2+\dots \text{ millions of determinants} \arrow[dr] & \\
& & \begin{vmatrix} \psi_1(\mathbf{x}_1) & \dots & \psi_1(\mathbf{x}_N) \\ \vdots & & \vdots \\ \psi_{N+1}(\mathbf{x}_1) & \dots & \psi_{N+1}(\mathbf{x}_N) \end{vmatrix}
\end{tikzcd}$$

with spin-orbitals $\psi_i(\mathbf{x})=\phi_i(\mathbf{r})\chi_{s_i}(\sigma)$, $s_i=\uparrow,\downarrow$

Variational principle: minimize $\langle\Psi|\mathcal{H}|\Psi\rangle$

Analytical integral → Gaussian basis

Quantum Monte Carlo wave function

Jastrow-Slater wave functions

$$\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)=\sum_k d_k D_k^{\uparrow} D_k^{\downarrow} \;\times\; \mathcal{J}$$

$$\swarrow \qquad\qquad \searrow$$

Determinantal part   Jastrow factor

$$\left.\begin{array}{l}\text{e-N}\\ \text{e-e}\\ \text{e-e-N}\end{array}\right\}\text{correlation}$$

- $\sum_k d_k D_k^{\uparrow} D_k^{\downarrow} \longrightarrow \boxed{\text{Few}}$ Slater determinants

  ↑, ↓-spin determinants of single-particle orbitals:

  - Slater functions for all-electron calculations

$$\phi(\mathbf{r})=\sum_{\alpha k} c_{k_\alpha}\, N_{k_\alpha} r_\alpha^{n_{k_\alpha}-1}\, e^{-\zeta_{k_\alpha} r_\alpha}\, \mathrm{Y}_{l_{k_\alpha} m_{k_\alpha}}(\hat{\mathrm{r}}_\alpha)$$

  - Gaussians for pseudopotential calculations

- $\mathcal{J} \longrightarrow$ Electron-electron correlation (e-e distance r_{ij})

Why can we factorize $D^{\uparrow}D^{\downarrow}$? (1)

Wave function in terms of space + spin variables:

$$\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N)=\Psi(\mathbf{r}_1,\sigma_1,\ldots,\mathbf{r}_N,\sigma_N) \quad \text{with} \quad \sigma_i=\pm 1$$

Consider $N=N_\uparrow+N_\downarrow$ and $S_z=(N_\uparrow-N_\downarrow)/2$ and

$$\zeta_1(\sigma_1,\ldots,\sigma_N)=\chi_\uparrow(\sigma_1)\ldots\chi_\uparrow(\sigma_{N_\uparrow})\chi_\downarrow(\sigma_{N_\uparrow+1})\ldots\chi_\downarrow(\sigma_N)$$

$\zeta_i(\sigma_1,\ldots,\sigma_N)$ generated by permuting indices in ζ_1

form a complete orthonormal set in spin space

$$\sum_{\sigma_1\ldots\sigma_N}\zeta_i(\sigma_1,\ldots,\sigma_N)\zeta_j(\sigma_1,\ldots,\sigma_N)=\delta_{ij}$$

$$\Rightarrow \quad \Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N)=\sum_{i=1}^{K}F_i(\mathbf{r}_1,\ldots,\mathbf{r}_N)\zeta_i(\sigma_1,\ldots,\sigma_N)$$

where F_i antisymmetric for interchange of like-spin

F_i equal to ± permutation of F_1

$$\boxed{\Psi(\mathbf{x}_1,\ldots,\mathbf{x}_N)=\mathcal{A}\left\{F_1(\mathbf{r}_1,\ldots,\mathbf{r}_N)\zeta_1(\sigma_1,\ldots,\sigma_N)\right\}}$$

Why can we factorize $D^{\uparrow}D^{\downarrow}$? (2)

Note that if $\mathcal{O}$ is a spin-independent operator

$$\boxed{\langle\Psi|\mathcal{O}|\Psi\rangle = \langle F_1|\mathcal{O}|F_1\rangle}$$

since ζ_i form an orthonormal set

More convenient to use F_1 instead of Ψ

To obtain F_1, assign the spin-variables of particles:

Particle	1	2	...	$N_\uparrow$	$N_{\uparrow+1}$	...	N
σ	1	1	...	1	-1	...	-1

$$\boxed{F_1(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \Psi(\mathbf{r}_1,1,\ldots,\mathbf{r}_{N_\uparrow},1,\mathbf{r}_{N_\uparrow+1},-1,\ldots,\mathbf{r}_N,-1)}$$

$\boxed{\text{Spin-assigned } \Psi = D}$

Determinant D of spin-orbitals $\psi_i(\mathbf{x}) = \phi_i(\mathbf{r})\chi_{s_i}(\sigma)$

Example: Be atom, $1s^2\,2s^2 \Rightarrow N_\uparrow = N_\downarrow = 2$, $S_z = 0$

Spin-orbitals $\phi_{1s}\,\chi_\uparrow$, $\phi_{2s}\,\chi_\uparrow$, $\phi_{1s}\,\chi_\downarrow$, $\phi_{2s}\,\chi_\downarrow$

$$D = \frac{1}{\sqrt{4!}} \begin{vmatrix} \phi_{1s}(\mathbf{r}_1)\chi_\uparrow(\sigma_1) & \ldots & \phi_{1s}(\mathbf{r}_4)\chi_\uparrow(\sigma_4) \\ \phi_{2s}(\mathbf{r}_1)\chi_\uparrow(\sigma_1) & \ldots & \phi_{2s}(\mathbf{r}_4)\chi_\uparrow(\sigma_4) \\ \phi_{1s}(\mathbf{r}_1)\chi_\downarrow(\sigma_1) & \ldots & \phi_{1s}(\mathbf{r}_4)\chi_\downarrow(\sigma_4) \\ \phi_{2s}(\mathbf{r}_1)\chi_\downarrow(\sigma_1) & \ldots & \phi_{2s}(\mathbf{r}_4)\chi_\downarrow(\sigma_4) \end{vmatrix}$$

$$F_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = D(\mathbf{r}_1, +1, \mathbf{r}_2, +1, \mathbf{r}_3, -1, \mathbf{r}_4, -1)$$

$$F_1 = \frac{1}{\sqrt{4!}} \begin{vmatrix} \phi_{1s}(\mathbf{r}_1) & \phi_{1s}(\mathbf{r}_2) & 0 & 0 \\ \phi_{2s}(\mathbf{r}_1) & \phi_{2s}(\mathbf{r}_2) & 0 & 0 \\ 0 & 0 & \phi_{1s}(\mathbf{r}_3) & \phi_{1s}(\mathbf{r}_4) \\ 0 & 0 & \phi_{2s}(\mathbf{r}_3) & \phi_{2s}(\mathbf{r}_4) \end{vmatrix}$$

$$D \to \begin{vmatrix} \phi_{1s}(\mathbf{r}_1) & \phi_{1s}(\mathbf{r}_2) \\ \phi_{2s}(\mathbf{r}_1) & \phi_{2s}(\mathbf{r}_2) \end{vmatrix} \times \begin{vmatrix} \phi_{1s}(\mathbf{r}_3) & \phi_{1s}(\mathbf{r}_4) \\ \phi_{2s}(\mathbf{r}_3) & \phi_{2s}(\mathbf{r}_4) \end{vmatrix} = D^\uparrow \times D^\downarrow$$

Spin-assigned $\Psi = \sum_k d_k D_k$

Care with order of spin-orbitals in determinants

e.g. First all ↑ spin-orbitals, then all ↓ spin-orbitals

Example: He atom, singlet excited state $1s^1 2s^1$

Spin-orbitals $\phi_{1s}\chi_\uparrow$, $\phi_{1s}\chi_\downarrow$, $\phi_{2s}\chi_\uparrow$, $\phi_{2s}\chi_\downarrow$

$$\begin{aligned}
\Psi &= \begin{vmatrix} \phi_{1s}(\mathbf{r_1})\chi_\uparrow(\sigma_1) & \phi_{1s}(\mathbf{r_2})\chi_\uparrow(\sigma_2) \\ \phi_{2s}(\mathbf{r_1})\chi_\downarrow(\sigma_1) & \phi_{2s}(\mathbf{r_2})\chi_\downarrow(\sigma_2) \end{vmatrix} \\
&- \begin{vmatrix} \phi_{1s}(\mathbf{r_1})\chi_\downarrow(\sigma_1) & \phi_{1s}(\mathbf{r_2})\chi_\downarrow(\sigma_2) \\ \phi_{2s}(\mathbf{r_1})\chi_\uparrow(\sigma_1) & \phi_{2s}(\mathbf{r_2})\chi_\uparrow(\sigma_2) \end{vmatrix} \\
&= \begin{vmatrix} \phi_{1s}(\mathbf{r_1})\chi_\uparrow(\sigma_1) & \phi_{1s}(\mathbf{r_2})\chi_\uparrow(\sigma_2) \\ \phi_{2s}(\mathbf{r_1})\chi_\downarrow(\sigma_1) & \phi_{2s}(\mathbf{r_2})\chi_\downarrow(\sigma_2) \end{vmatrix} \\
&+ \begin{vmatrix} \phi_{2s}(\mathbf{r_1})\chi_\uparrow(\sigma_1) & \phi_{2s}(\mathbf{r_2})\chi_\uparrow(\sigma_2) \\ \phi_{1s}(\mathbf{r_1})\chi_\downarrow(\sigma_1) & \phi_{1s}(\mathbf{r_2})\chi_\downarrow(\sigma_2) \end{vmatrix}
\end{aligned}$$

Assign spins:	Particle	1	2
	σ	1	-1

$$\boxed{F_1(\mathbf{r_1}, \mathbf{r_2}) = \phi_{1s}(\mathbf{r_1})\phi_{2s}(\mathbf{r_2}) + \phi_{2s}(\mathbf{r_1})\phi_{1s}(\mathbf{r_2})}$$

Spin-assigned QMC wave functions

$\sigma = +1$ for first $N_\uparrow$ particles, $\sigma = -1$ for the others

$$\begin{aligned}\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N) &= F_1(\mathbf{r}_1,\ldots,\mathbf{r}_N) \\ &= \mathcal{J}\sum_k d_k\, D_k^\uparrow(\mathbf{r}_1,\ldots,\mathbf{r}_{N_\uparrow}) D_k^\downarrow(\mathbf{r}_{N_\uparrow+1},\ldots,\mathbf{r}_N)\end{aligned}$$

where $\mathcal{J} = \mathcal{J}(\mathbf{r}_1,\ldots,\mathbf{r}_N)$ is the Jastrow factor

Spacial symmetry

$\sum_k d_k\, D_k$ constructed to have proper spacial symmetry

Often, $\mathcal{J} = \mathcal{J}(\{r_{ij}\},\{r_{i\alpha}\})$, i,j=electrons, α=nucleus

$\Rightarrow$ $\mathcal{J}$ invariant under rotations

$\Rightarrow$ $\mathcal{J}$ does not affect spacial symmetry of Ψ

Spin symmetry

$\sum_k d_k\, D_k$ constructed to be eigenstate of S^2, S_z

$\mathcal{J}$ symmetric for interchange of like-spin particles

$\Rightarrow$ Ψ eigenstate of S_z

$\mathcal{J}$ symmetric for interchange of spacial variables

$\Rightarrow$ Ψ eigenstate of S^2

Cusp conditions

At interparticle coalescence points, potential diverges:

Electron-nucleus $-\dfrac{Z}{r_{i\alpha}}$

Electron-electron $\dfrac{1}{r_{ij}}$

Local energy $\boxed{\dfrac{\mathcal{H}\Psi}{\Psi} = -\dfrac{1}{2}\sum_i \dfrac{\nabla_i^2\Psi}{\Psi} + V}$ must be finite

⇒ Kinetic energy must have opposite divergence

⇒ Ψ must satisfy Kato's cusp conditions:

$$\boxed{\left.\frac{\partial\hat{\Psi}}{\partial r_{ij}}\right|_{r_{ij}=0} = \mu_{ij} q_i\, q_j \Psi(r_{ij}=0)}$$

for two particles of masses m_i, m_j and charges q_i, q_j

Note: all other interparticle distances are > 0,

$\hat{\Psi}$ is a spherical average, and $\mu_{ij} = \dfrac{m_i\, m_j}{m_i + m_j}$

Cusp conditions: example

Consider $r_{ij} \to 0$ and all other particles well separated

The local energy close to $r = r_{ij} = 0$ is:

$$-\frac{1}{2\mu_{ij}}\frac{\nabla^2\Psi}{\Psi} + V(r) \;=\; \text{finite}$$

Assume $\boxed{\Psi(r = r_{ij} = 0) \neq 0}$

$$-\frac{1}{2\mu_{ij}}\frac{\Psi''}{\Psi} - \frac{1}{\mu_{ij}}\frac{1}{r}\frac{\Psi'}{\Psi} + V(r) \;=\; \text{finite}$$

The condition for E_L to be finite at $r = 0$ is

$$\frac{\Psi'}{\Psi} = \mu_{ij}\, r\, V(r)$$

- Electron-nucleus: $V = -\frac{Z}{r}, \mu = 1 \Rightarrow \boxed{\left.\frac{\Psi'}{\Psi}\right|_{r=0} = -Z}$
- Electron-electron: $V = \frac{1}{r}, \quad \mu = \frac{1}{2} \Rightarrow \boxed{\left.\frac{\Psi'}{\Psi}\right|_{r=0} = 1/2}$

Generalized cusp conditions

R. T. Pack and W. Byers Brown, J. Chem. Phys. **45**, 556 (1966)

What about two electrons in a triplet state?

Or more generally two like-spin electrons ($D \to 0$)?

Or a highly excited state (e.g. $2p^2$ state of Helium)?

$$\boxed{\Psi(r = r_{ij} = 0) = 0} \quad ?!?$$

Wave function near $\mathbf{r} = \mathbf{r}_{ij} = 0$ can be written as:

$$\Psi = \sum_{l=l_0}^{\infty} \sum_{m=-l}^{l} f_{lm}(r)\, r^l\, Y_{lm}(\theta, \phi)$$

Expanding $f_{lm}(r) = \sum_{k=0}^{\infty} f_{lm}^{(k)}\, r^k$

$$\boxed{f_{lm}(r) = f_{lm}^{(0)} \left[1 + \frac{\gamma}{(l+1)} r + O(r^2)\right]}$$

where $\gamma = q_i\, q_j\, \mu_{ij}$

- Electron-electron singlet: $l_0 = 0 \Rightarrow \boxed{\Psi \sim \left(1 + \frac{1}{2} r\right)}$
- Electron-electron triplet: $l_0 = 1 \Rightarrow \boxed{\Psi \sim \left(1 + \frac{1}{4} r\right)\, r}$

Cusp conditions and QMC wave functions (1)

$\sigma = +1$ for first $N_\uparrow$ electrons, $\sigma = -1$ for the others

$$\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \mathcal{J} \sum_k d_k\, D_k^\uparrow(\mathbf{r}_1,\ldots,\mathbf{r}_{N_\uparrow}) D_k^\downarrow(\mathbf{r}_{N_\uparrow+1},\ldots,\mathbf{r}_N)$$

Electron-electron cusp conditions

- Anti-parallel spins: $r_{ij} \to 0$ for $i \le N_\uparrow$, $j \ge N_\uparrow + 1$

  Usually, determinantal part $\neq 0$

  $$\Rightarrow \quad \mathcal{J}(r_{ij}) \sim \left(1 + \frac{1}{2} r_{ij}\right) \quad \Leftrightarrow \quad \boxed{\left.\frac{\mathcal{J}'}{\mathcal{J}}\right|_{r_{ij}=0} = \frac{1}{2}}$$

- Parallel spins: $r_{ij} \to 0$ for $i, j \le N_\uparrow$ or $i, j \ge N_\uparrow + 1$

  Determinantal part $\to 0$

  $$\Rightarrow \quad \mathcal{J}(r_{ij}) \sim \left(1 + \frac{1}{4} r_{ij}\right) \quad \Leftrightarrow \quad \boxed{\left.\frac{\mathcal{J}'}{\mathcal{J}}\right|_{r_{ij}=0} = \frac{1}{4}}$$

$\Rightarrow$ $\mathcal{J}$ not symmetric for interchange of $\mathbf{r}_1,\ldots,\mathbf{r}_N$

$\Rightarrow$ Ψ is not an eigenstate of S^2

For optimized Ψ, spin contamination is small

Huang, Filippi, Umrigar, J. Chem. Phys. **108**, 8838 (1998)

Cusp conditions and QMC wave functions (2)

$\sigma = +1$ for first $N_\uparrow$ electrons, $\sigma = -1$ for the others

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \mathcal{J} \sum_k d_k \, D_k^\uparrow(\mathbf{r}_1, \ldots, \mathbf{r}_{N_\uparrow}) D_k^\downarrow(\mathbf{r}_{N_\uparrow+1}, \ldots, \mathbf{r}_N)$$

Electron-nucleus cusp conditions

Usually, imposed through the determinantal part

Assume nucleus at the origin and $\Psi(r_i = 0) \neq 0$

If each orbital satisfies the cusp conditions

$$\begin{aligned} \left.\frac{\partial \widehat{\phi}_j}{\partial r}\right|_{r=0} &= -Z\widehat{\phi}_j(r=0) \\ \Rightarrow \quad \left.\frac{\partial \sum_k d_k \, \widehat{D}_k}{\partial r}\right|_{r=0} &= -Z \sum_k d_k \, \widehat{D}_k(r=0) \end{aligned}$$

Note: Slater basis best suited for all-electron systems

No electron-nucleus cusp with pseudopotential

Cusp conditions in Be atom

Be atom, $1s^2\,2s^2 \Rightarrow N_\uparrow = N_\downarrow = 2,\ S_z = 0$

Spin-assigned $\Psi(\mathbf{r}_1^+, \mathbf{r}_2^+, \mathbf{r}_3^-, \mathbf{r}_4^-) = \mathcal{J}\, D$

- Factorized determinant

$$D = D^\uparrow \times D^\downarrow = \begin{vmatrix} \phi_{1s}(\mathbf{r}_1) & \phi_{1s}(\mathbf{r}_2) \\ \phi_{2s}(\mathbf{r}_1) & \phi_{2s}(\mathbf{r}_2) \end{vmatrix} \times \begin{vmatrix} \phi_{1s}(\mathbf{r}_3) & \phi_{1s}(\mathbf{r}_4) \\ \phi_{2s}(\mathbf{r}_3) & \phi_{2s}(\mathbf{r}_4) \end{vmatrix}$$

If $\phi_{1s}(\mathbf{r}) = c_1\, \mathrm{e}^{-\zeta_1 r} + c_2\, \mathrm{e}^{-\zeta_2 r} + c_3\, r\, \mathrm{e}^{-\zeta_3 r} + c_4\, r\, \mathrm{e}^{-\zeta_4 r}$

$$\left.\frac{\partial \phi_{1s}}{\partial r}\right|_{r=0} = -Z\phi_{1s}(0) \Rightarrow c_1 = \frac{c_2(Z-\zeta_2) + c_3 + c_4}{\zeta_1 - Z}$$

- Simple Jastrow factor

$$\mathcal{J} = \prod_{ij=13,14,23,24} \exp\left\{\frac{1}{2}\frac{r_{ij}}{1+b\, r_{ij}}\right\} \times$$

$$\times \prod_{ij=12,34} \exp\left\{\frac{1}{4}\frac{r_{ij}}{1+b\, r_{ij}}\right\}$$

Jastrow factor for atoms and molecules

Simple Jastrow factor

$$\mathcal{J}(r_{ij}) = \prod_{i<j} \exp\left\{ b_0 \frac{r_{ij}}{1 + b\, r_{ij}} \right\} \quad \text{with} \quad b_0 = \frac{1}{2} \text{ or } \frac{1}{4}$$

Boys and Handy's form

$$\mathcal{J}(r_i, r_j, r_{ij}) = \prod_{\alpha, i<j} \exp\left\{ \sum c^{\alpha}_{mnk} \left(\bar{r}^{m}_{i\alpha}\, \bar{r}^{n}_{j\alpha} + \bar{r}^{n}_{i\alpha}\, \bar{r}^{m}_{j\alpha} \right) \bar{r}^{k}_{ij} \right\}$$

with $\bar{r}_{i\alpha} = \dfrac{a\, r_{i\alpha}}{1 + a\, r_{i\alpha}}$ and $\bar{r}_{ij} = \dfrac{d\, r_{ij}}{1 + d\, r_{ij}}$

Cusp conditions imposed by requiring:

$m = n = 0$ if $k = 1$ for electron-electron cusps

No $n = 1$ or $m = 1$, D satisfies electron-nucleus cusps

More general form

Lift constraints and allow all values of n, m, k

Cusp conditions ⇒ linear dependencies among c^{α}_{mnk}

Other scaling functions are possible: $(1 - \mathrm{e}^{-a\, r})/a \ldots$

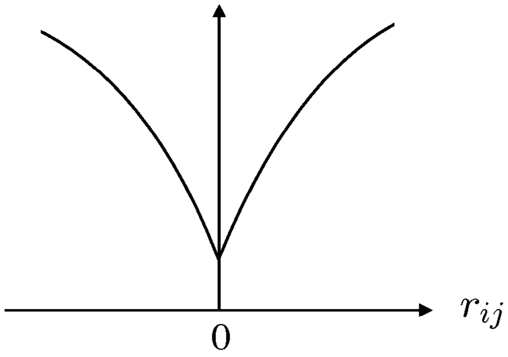
Some comments on Jastrow factor

- $\mathcal{J} > 0$ and becomes constant for large r_i, r_j and r_{ij} (ratio of polynomials or use of scaled variables)
- Preferable to separate e-n, e-e and e-e-n terms as

$$\prod_{\alpha,i} \exp\left\{A(r_{i\alpha})\right\} \prod_{i<j} \exp\left\{B(r_{ij})\right\} \prod_{\alpha,i<j} \exp\left\{C(r_{i\alpha}, r_{j\alpha}, r_{ij})\right\}$$

Electron-electron terms

- Introduced to impose the cusp conditions and to keep electrons apart, e.g. simple $\mathcal{J}(r_{ij})$ looks like



- No significant improvement in using $\mathcal{J}(r_{ij})$ more general than simple $\mathcal{J}$ but not a function of r_i, r_j

Electron-nucleus terms

- Omissible if the determinantal part is constructed with a sufficiently large basis and then reoptimized

- The e-n terms should be included if the determinantal part (often DFT or HF) is not reoptimized: the e-e terms alter the single-particle density (reduced/increased in high/low density regions)

  Electron-electron-nucleus terms

- If the order of the polynomial in the e-e-n terms is infinite, the wave function can exactly describe a two-electron atom or ion in an S state

  For these systems, a 5th-order polynomial recovers more than 99.99% of the correlation energy

- Is this wave function adequate for multi-electron systems? The e-e-n terms are the most important ones: due to the exclusion principle, it is rare for 3 or more electrons to be close, since at least 2 electrons must necessarily have the same spin

- Ratio of polynomials or higher-order polynomial? For instance, for 1st-row diatomics, ratio of two 4th-order polynomials and a 5th-order polynomial about the same quality

Jastrow factor with e-e, e-e-n and e-e-e-n terms

	$\mathcal{J}$	E_{VMC}	$E^{\mathrm{corr}}_{\mathrm{VMC}}$ (%)	σ_{VMC}
Li				
E_{HF}		-7.43273	0	
	e-e	-7.47427(4)	91.6	0.24
	+ e-e-n	-7.47788(1)	99.6	0.037
	+ e-e-e-n	-7.47797(1)	99.8	0.028
E_{exact}		-7.47806	100	
Be				
E_{HF}		-14.57302	0	
	e-e	-14.66088(5)	93.1	0.35
	+ e-e-n	-14.66662(1)	99.2	0.089
	+ e-e-e-n	-14.66681(1)	99.4	0.078
E_{exact}		-14.66736	100	
Ne				
E_{HF}		-128.5471	0	
	e-e	-128.713(2)	42.5	1.9
	+ e-e-n	-128.9008(1)	90.6	0.90
	+ e-e-e-n	-128.9029(3)	91.1	0.88
E_{exact}		-128.9376	100	

Huang, Umrigar, Nightingale, J. Chem. Phys. **107**, 3007 (1997)

Dynamic and static correlation

There are two types of correlation:

- Dynamic correlation

  Due to inter-electron repulsion and always present

  Described by Jastrow factor

- Static correlation

  Due to near-degeneracy of occupied and unoccupied orbitals and not always present

  Described by a linear combination of determinants

Example: Be atom and $2s$-$2p$ near-degeneracy

HF ground state configuration	$1s^2 2s^2$
Additional important configuration	$1s^2 2p^2$

Ground state has ${}^1S$ symmetry ⇒ 4 determinants:

$$
\begin{aligned}
D = (1s^\uparrow, 2s^\uparrow, 1s^\downarrow, 2s^\downarrow) + c * \Big\{ & (1s^\uparrow, 2p_x^\uparrow, 1s^\downarrow, 2p_x^\downarrow) \\
+ & (1s^\uparrow, 2p_y^\uparrow, 1s^\downarrow, 2p_y^\downarrow) \\
+ & (1s^\uparrow, 2p_z^\uparrow, 1s^\downarrow, 2p_z^\downarrow) \Big\}
\end{aligned}
$$

$$
\boxed{
\begin{array}{lcl}
1s^2 2s^2 & \times \mathcal{J}(r_{ij}) & \rightarrow E_{\text{VMC}}^{\text{corr}} = 61\% \\
1s^2 2s^2 \oplus 1s^2 2p^2 & \times \mathcal{J}(r_{ij}) & \rightarrow E_{\text{VMC}}^{\text{corr}} = 93\%
\end{array}
}
$$

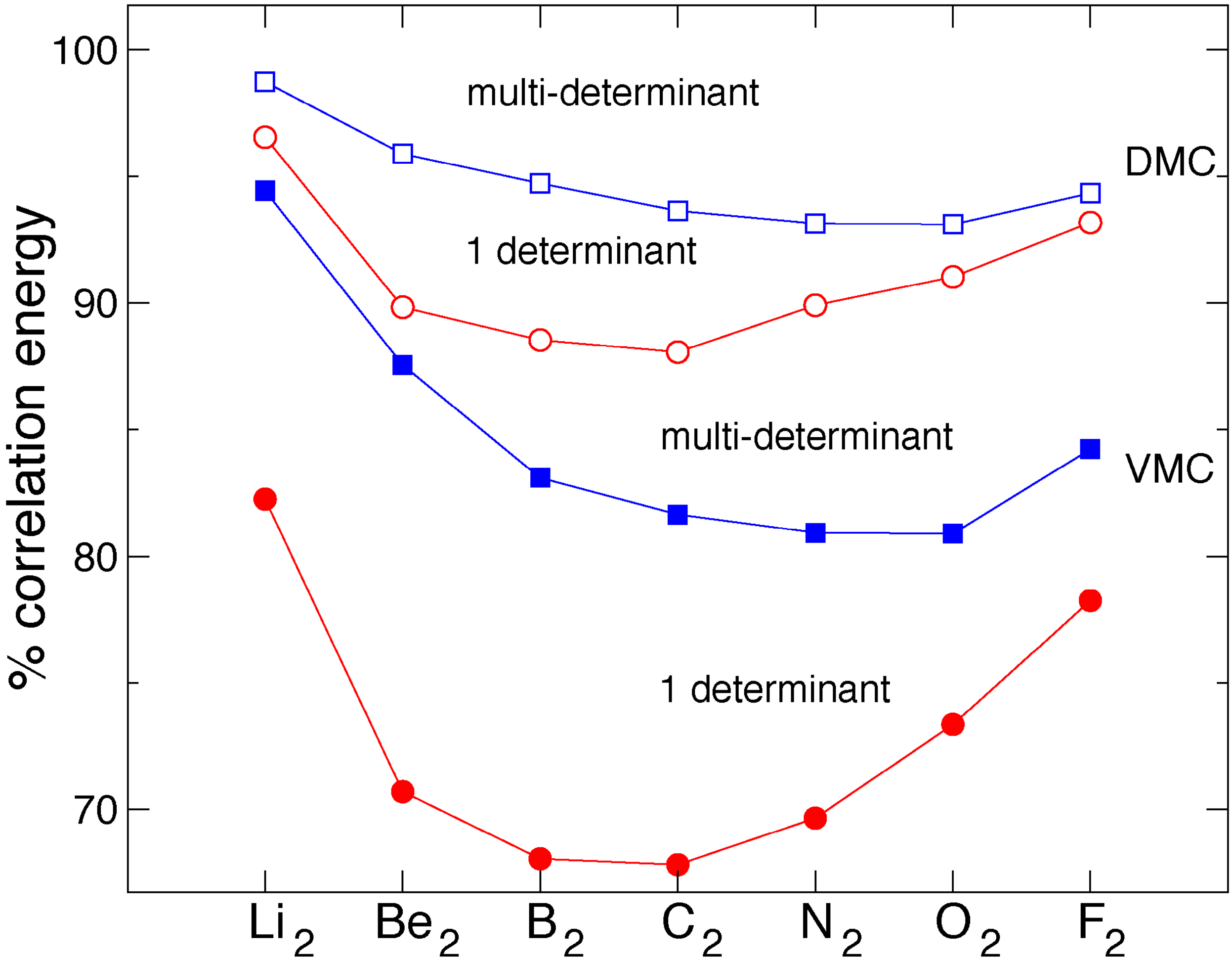
Static correlation

Example: 1^{st}-row dimers (all-electron calculations)

MO orbitals with atomic s-p Slater basis

Active MO's: $2\sigma_g, 2\sigma_u, 3\sigma_g, 3\sigma_u, 1\pi_u, 1\pi_g$

5^{th}-order polynomial $\mathcal{J}$ (e-n, e-e, e-e-n)



Filippi and Umrigar, J. Chem. Phys. **105**, 213 (1996)

Determinant versus Jastrow factor

Determinantal part yields the nodes of wave function

⇒ Quality of the fixed-node DMC solution

(see tomorrow diffusion Monte Carlo)

Why bother with the Jastrow factor?

Implications of using a good Jastrow factor for DMC

- Efficiency: Smaller $\sigma \Rightarrow$ gain in CPU time
  (also smaller time-step error)
- Expectation values other than energy
  Mixed estimator
- Pseudopotentials: Localization error
  Jastrow factor does affect fixed-node energy

Why should $\Psi = \mathcal{J}D$ work?

$$\begin{tikzcd}
\begin{array}{c}\text{Full wave function}\\ \Psi\end{array} \arrow[r] \arrow[d] & \begin{array}{c}\text{Factorized wave function}\\ \mathcal{J}\Phi\end{array} \arrow[d] \\
\begin{array}{c}\text{Full Hamiltonian}\\ \mathcal{H}\end{array} \arrow[r] & \begin{array}{c}\text{Effective Hamiltonian}\\ \mathcal{H}_{\text{eff}}\end{array} \\
\boxed{\mathcal{H}\Psi = \text{E}\Psi} \arrow[r] & \mathcal{H}\mathcal{J}\Phi = \text{E}\mathcal{J}\Phi
\end{tikzcd}$$

$$\frac{\mathcal{H}\mathcal{J}}{\mathcal{J}}\Phi = \text{E}\Phi$$

$$\boxed{\mathcal{H}_{\text{eff}}\Phi = \text{E}\Phi}$$

$\mathcal{H}_{\text{eff}}$ weaker Hamiltonian than $\mathcal{H}$

$\Rightarrow$ $\Phi \approx$ non-interacting wave function D

$\Rightarrow$ Quantum Monte Carlo wave function $\Psi = \mathcal{J}D$

Optimization of trial wave function

Start from $\Psi_T(\mathbf{R}\,,\{\alpha_0\})$ with parameters $\{\alpha_0\}$

Generate N_{conf} walkers distributed as $|\Psi_T(\mathbf{R}\,,\{\alpha_0\})|^2$

How do we find a better set of parameters $\{\alpha\}$?

First thought: Minimize the energy

$$E[\alpha] = \frac{1}{N_{\text{conf}}} \sum_{i=1}^{N_{\text{conf}}} \frac{\mathcal{H}\Psi(\mathbf{R}_i\,,\{\alpha\})}{\Psi(\mathbf{R}_i\,,\{\alpha\})} = \frac{1}{N_{\text{conf}}} \sum_{i=1}^{N_{\text{conf}}} E_{\text{L}}(\mathbf{R}_i,\{\alpha\})$$

Straightforward minimization of $E[\alpha]$ does not work:

N_{conf} is a relatively small number of configurations

$\Rightarrow$ $E[\alpha]$ unbounded from below

$\Rightarrow$ Usually, one finds lower $E[\alpha]$ on the given set of $\mathbf{R}_i$

but a higher energy in a new VMC run

Better method: Minimize variance of local energy

Coldwell, Int. J. Quantum Chem. Symp. **11**, 215 (1977)

Umrigar, Wilson, Wilkins, Phys. Rev. Lett. **60**, 1719 (1988)

Variance minimization (1)

Generate $N_{\rm conf}$ walkers distributed as $|\Psi(\mathbf{R}, \{\alpha_0\})|^2$

Minimize the variance of the local energy $\sigma^2[\alpha]$:

$$\sigma^2[\alpha] = \sum_{i=1}^{N_{\rm conf}} \left(\frac{\mathcal{H}\Psi(\mathbf{R}_i, \{\alpha\})}{\Psi(\mathbf{R}_i, \{\alpha\})} - \bar{E} \right)^2 w_i$$

where

$$w_i = \left| \frac{\Psi(\mathbf{R}_i, \{\alpha\})}{\Psi(\mathbf{R}_i, \{\alpha_0\})} \right|^2 \Bigg/ \sum_{i=1}^{N_{\rm conf}} \left| \frac{\Psi(\mathbf{R}_i, \{\alpha\})}{\Psi(\mathbf{R}_i, \{\alpha_0\})} \right|^2$$

and $\bar{E}$ is the average energy

- Why do we introduce the weights w_i?

1) To provide correct reweight as Ψ changes

2) To allow nodes to move during optimization

Note: w_i needs to be limited to a maximum value

(few $\mathbf{R}_i$ may gain large w_i and dominate minimization)

- $\bar{E}$ substituted with $E_{\rm guess}$

$E_{\rm guess}$ chosen a bit less than current energy estimate

$\Leftrightarrow$ Minimize a combination of variance and energy

Variance minimization

Some advantages:

- σ^2 has a known lower bound: $\sigma^2 = 0$
- All eigenstates have zero variance

  $\Rightarrow$ It is possible to optimize excited states (also a higher lying state of a given symmetry)
- Cusp conditions or other constraints easily added

  $\Rightarrow$ Minimize $\quad \chi^2 = \sigma^2 +$ penalty functions
- Efficient procedures to optimize a sum of squares:
  - It is helpful to know not only the gradient but also the Hessian of the quantity being optimized. If one minimizes a sum of squares, it is possible to calculate an approximate second derivative matrix using only the first derivatives
  - Efficient methods, e.g. Levenberg-Marquard
- N_{conf} =2000-3000 sufficient for 50-100 parameters for large dimensional spaces ($\sim$ 800 dim)

Variance minimization

Why do we need so few N_{conf}?

- In the optimization, the configurations are fixed
  $\Leftrightarrow$ Correlated sampling: The difference $\sigma[\{\alpha\}]^2 - \sigma[\{\alpha_0\}]^2$ is better determined than separate σ's
- We are performing a fit not an integral

Some disadvantages:

- It is variance not energy minimization!

  For a given functional form, different parameter sets $\{\alpha\}$ can give comparable σ but different E_{VMC} (in particular if one optimizes determinantal part)
- It is a non-linear optimization

  $\Rightarrow$ It is possible to get stuck in local minima
- Easy optimization of the Jastrow factor

  More tricky for the determinantal component

Variance minimization

Operationally

1. Start from initial wave function $\Psi(\mathbf{R},\{\alpha_0\})$, e.g.
   - HF or MCSCF-determinant + simple Jastrow (set b to a reasonable value, $b \approx 0.5 - 1$)
   - For simple systems, guess LCAO and basis exponents. With some experience, it will work!
2. Do a VMC run to sample $|\Psi(\mathbf{R},\{\alpha_0\})|^2$
   - Generate N_{conf} walkers $\{\mathbf{R}_i\}$
   - Set E_{guess} a bit lower than E_{VMC}

     At 1st iteration, try $E_{\text{guess}} \approx E_{\text{VMC}} - 0.1 * \sigma_{\text{VMC}}$

     or use $E_{\text{corr}} \approx 0.4 - 1.2$ eV/elec for $Z < 18$
3. Optimize $\sigma^2[\alpha] \Rightarrow$ new set of parameters $\{\alpha_1\}$
4. Do a VMC run to sample $|\Psi(\mathbf{R},\{\alpha_1\})|^2$

   a) If $E_{\text{VMC}}[\{\alpha_1\}]$ is lower than $E_{\text{VMC}}[\{\alpha_0\}]$
   - Generate new N_{conf} walkers $\{\mathbf{R}_i\}$
   - Set E_{guess} to new E_{VMC}
   - Iterate 3-4

Continue . . .

b) If $E_{\mathsf{VMC}}[\{\alpha_1\}]$ is higher than $E_{\mathsf{VMC}}[\{\alpha_0\}]$

- Bad starting wave function?
- Too many parameters varied at once?
- E_{guess} too low?

$\Rightarrow$ Do not update parameters

- Go back to step 1 and/or 3

5. Perform long VMC run with optimal final $\{\alpha\}$

Quality of wave function $\Leftrightarrow$ E_{VMC} and σ_{VMC}

What about diffusion Monte Carlo?

$\Psi(\mathbf{R}, \{\alpha\}) \Rightarrow$ improved σ_{DMC} and (usually) E_{DMC}

Note: E_{guess} usually converged in 2 iterations

For simplicity, one can set $w_i = 1$ in σ^2

⋆ **This afternoon**

Optimization by variance minimization

- **Be atom**

  $1s^2 2s^2$ + simple Jastrow factor

  1. Vary b parameter in Jastrow factor
  2. Vary LCAO in Slater basis of $1s$, $2s$ orbitals

     Number of degrees of freedom in $1s$?

     N^{1s}_{LCAO} -1 (cusp), -1 (norm), -1 (pivot)

     Number of degrees of freedom in $2s$?

     N^{2s}_{LCAO} -1 (cusp), -1 (norm), -1 (pivot)
  3. Vary exponents of Slater basis

  $1s^2 2s^2 \oplus 1s^2 2p^2$ + simple Jastrow factor

  1. Start from 1-det wave function
  2. Vary b parameter in Jastrow factor
  3. Vary coefficient in front of $1s^2 2p^2$
  4. Vary LCAO in Slater basis of $1s$, $2s$, $2p$ orbitals

     Number of degrees of freedom in $1s$?

     N^{1s}_{LCAO} -1 (cusp), -1 (norm)

     Number of degrees of freedom in $2s$?

N^{2s}_{LCAO} -1 (cusp), -1 (norm), -1 (pivot)

Number of degrees of freedom in $2p$?

N^{2p}_{LCAO} -1 (norm)

5. Vary exponents of Slater basis

   Note: Relationship among p_x, p_y and p_z

- Homonuclear diatomic molecule B$_2$

$1\sigma_g^2\, 1\sigma_u^2\, 2\sigma_g^2\, 2\sigma_u^2\, 1\pi_{ux}\, 1\pi_{uy}$ + simple Jastrow factor

1. Vary b parameter in Jastrow factor
2. Vary LCAO in Slater basis on nuclei A and B

   Care with symmetry ⇒ linear dependencies

   $\sigma_g = c_1 * s^A + c_2 * p_z^A + c_1 * s^B - c_2 * p_z^B$

   $\sigma_u = c_1 * s^A + c_2 * p_z^A - c_1 * s^B + c_2 * p_z^B$

   $\pi_u = c_1 * p_x^A + c_1 * p_x^B$

   $\pi_g = c_1 * p_x^A - c_1 * p_x^B$

   Cusp conditions only on σ orbitals

   Pivoting among orbitals of same symmetry
3. Vary exponents of Slater basis

Multi-determinant + simple Jastrow factor

Effect of d-basis?

Energy minimization?

Subject of on-going research:

- Energy fluctuation potential method
- Stochastic reconfiguration
- Computation of derivatives and Hessian

...

Energy fluctuation potential method

Fahy, Filippi, Schautz, Prendergast, see references

Consider infinitesimal variations of Ψ

$$\Psi = \mathcal{J}\Phi \to \Psi' = \Psi + \sum_{k>0} \delta_k \frac{\partial \Psi}{\partial \alpha_k} = \Psi \left(1 + \sum_{k>0} \delta_k O_k \right)$$

with $O_k = \dfrac{1}{\Psi} \dfrac{\partial \Psi}{\partial \alpha_k}$

The energy is stationary if these derivatives are zero

$$\left.\frac{\partial E}{\partial \delta_k}\right|_{\delta=0} = \left.\frac{\partial}{\partial \delta_k} \frac{\langle \Psi' | \mathcal{H} | \Psi' \rangle}{\langle \Psi' | \Psi' \rangle}\right|_{\delta=0}$$

$$= \boxed{\langle (E_\mathrm{L} - \bar{E})\,(O_k - \bar{O}_k) \rangle_{\Psi^2}}$$

Energy fluctuation potential method

Energy stationary if $\langle (E_{\mathrm{L}} - \bar{E})\,(O_k - \bar{O}_k)\rangle_{\Psi^2} = 0$

$\Leftrightarrow$ The fluctuations of E_{L} and O_k are uncorrelated

$\Leftrightarrow$ E_{L} cannot be made 'more constant' by adding some combination of the functions O_k

Reformulate problem as a least-squares fit of E_{L}

$$\chi^2 = \langle\, (E_{\mathrm{L}} - E_0 - \sum_{k>0} V_k O_k)^2\,\rangle_{\Psi^2}$$

Equivalently, solve set of linear equations

$$\langle (E_{\mathrm{L}} - \bar{E})(O_m - \bar{O}_m)\rangle_{\Psi^2} = \sum_{k>0} V_k \langle (O_k - \bar{O}_k)(O_m - \bar{O}_m)\rangle_{\Psi^2}$$

Energy stationary $\Leftrightarrow$ $V_k = 0$

How do we use $V_k \neq 0$?

If one optimizes the determinantal part Φ, interpret fitting the fluctuations of $E_{\mathrm{L}} = \dfrac{\mathcal{H}\Psi}{\Psi}$ with $\sum_k V_k O_k$ as fitting the fluctuations of $\dfrac{\mathcal{H}_{\mathrm{eff}}\Phi}{\Phi}$ where $\mathcal{H}_{\mathrm{eff}} = \dfrac{\mathcal{H}\mathcal{J}}{\mathcal{J}}$

$\Rightarrow$ Use V_k to construct $\mathcal{H}'_{\mathrm{eff}}$ which approximates $\dfrac{\mathcal{H}\mathcal{J}}{\mathcal{J}}$ and use the solution of $\mathcal{H}'_{\mathrm{eff}}$ as new Φ

Energy fluctuation potential method

Example: Ground state of CH_2O, 1-det wave function

Pseudopotentials → 12 electrons

Optimization of orbitals with EFP method:

→ optimization of 315 LCAO parameters

	VMC	DMC
RHF orbitals	-22.763(3)	-22.8454(6)
Optimized	-22.784(2)	-22.8494(6)

Customary practice for optimizing wave function

Constructing wave function is a bit of an $\boxed{\text{art}}$

Jastrow-Slater wave function

$$\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N) = \underbrace{\sum_k d_k D_k^{\uparrow} D_k^{\downarrow}}_{\text{Determinantal part}} \times \underbrace{\mathcal{J}}_{\text{Jastrow factor}}$$

$$\left.\begin{array}{l}\text{e-N}\\ \text{e-e}\\ \text{e-e-N}\end{array}\right\}\ \text{correlation}$$

- Jastrow factor optimized in variance minimization
- Orbitals + d_k coefficients in determinantal part from

1. Hartree-Fock
2. Density functional theory (LDA, GGA, . . .)
3. Multi-configuration self-consistent-field
4. Optimized in variance minimization

   d_k coefficients (easy) + orbitals (small systems)
5. Energy minimization (active subject of research)

Correlated sampling in VMC

Two operators $\mathcal{O}$, $\mathcal{O}'$ and two wave functions Ψ, Ψ'

$$\boxed{\bar{\mathcal{O}}' - \bar{\mathcal{O}} = \frac{\langle\Psi'|\mathcal{O}'|\Psi'\rangle}{\langle\Psi'|\Psi'\rangle} - \frac{\langle\Psi|\mathcal{O}|\Psi\rangle}{\langle\Psi|\Psi\rangle}}$$

Correlated sampling is a technique to calculate differences more accurately than the separate quantities

Example: Map out potential energy surface

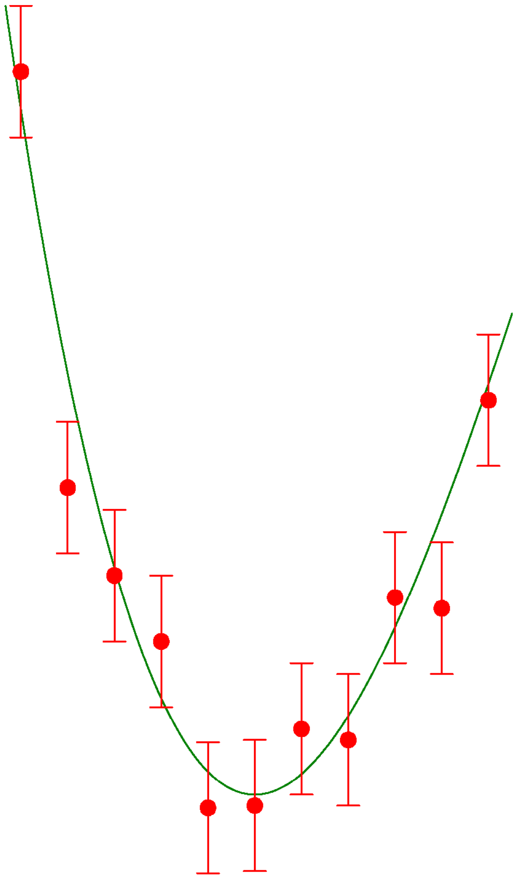
$\Rightarrow$ $\boxed{\text{Compute energy differences}}$

DFT/QC methods ⇒ smoothly varying error

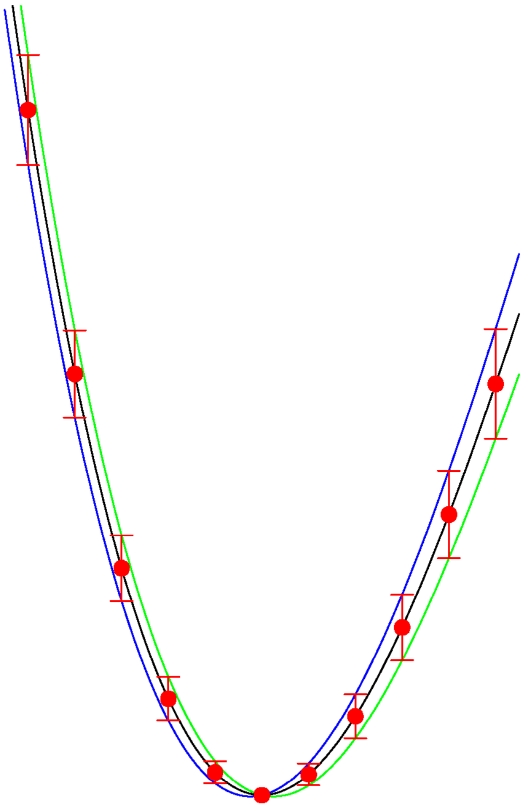
Problem in QMC: statistical error

Energy of a dimer versus bond length

Independent Monte Carlo runs



Correlated sampling



⇒ Forces cannot be computed from independent runs

Correlated sampling: e.g. potential energy surface

Primary geometry	$\mathcal{H}$	Ψ	E
Secondary geometry	$\mathcal{H}_{\mathrm{s}}$	Ψ_{s}	E_{s}

$$E_{\mathrm{s}} - E = \frac{\langle \Psi_{\mathrm{s}} | \mathcal{H}_{\mathrm{s}} | \Psi_{\mathrm{s}} \rangle}{\langle \Psi_{\mathrm{s}} | \Psi_{\mathrm{s}} \rangle} - \frac{\langle \Psi | \mathcal{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$

Do NOT perform independent MC runs

Generate MC configurations only from Ψ^2
where Ψ is the reference situation

$$E_{\mathrm{s}} - E = \frac{1}{N_{\mathrm{conf}}} \sum_{i=1}^{N_{\mathrm{conf}}} \left\{ \frac{\mathcal{H}_{\mathrm{s}} \Psi_{\mathrm{s}}(\mathbf{R}_i)}{\Psi_{\mathrm{s}}(\mathbf{R}_i)} w_i - \frac{\mathcal{H}\Psi(\mathbf{R}_i)}{\Psi(\mathbf{R}_i)} \right\}$$

$$w_i = \frac{|\Psi_{\mathrm{s}}(\mathbf{R}_i)/\Psi(\mathbf{R}_i)|^2}{\frac{1}{N_{\mathrm{conf}}} \sum_{j=1}^{N_{\mathrm{conf}}} \left|\Psi_{\mathrm{s}}(\mathbf{R}_j)/\Psi(\mathbf{R}_j)\right|^2}$$

Efficient if
- w_i not too different from 1
- $\mathcal{H}$ and $\mathcal{H}_{\mathrm{s}}$ closely related

Efficiency gain from correlated sampling

Example: B_2, 1 determinant + simple Jastrow factor

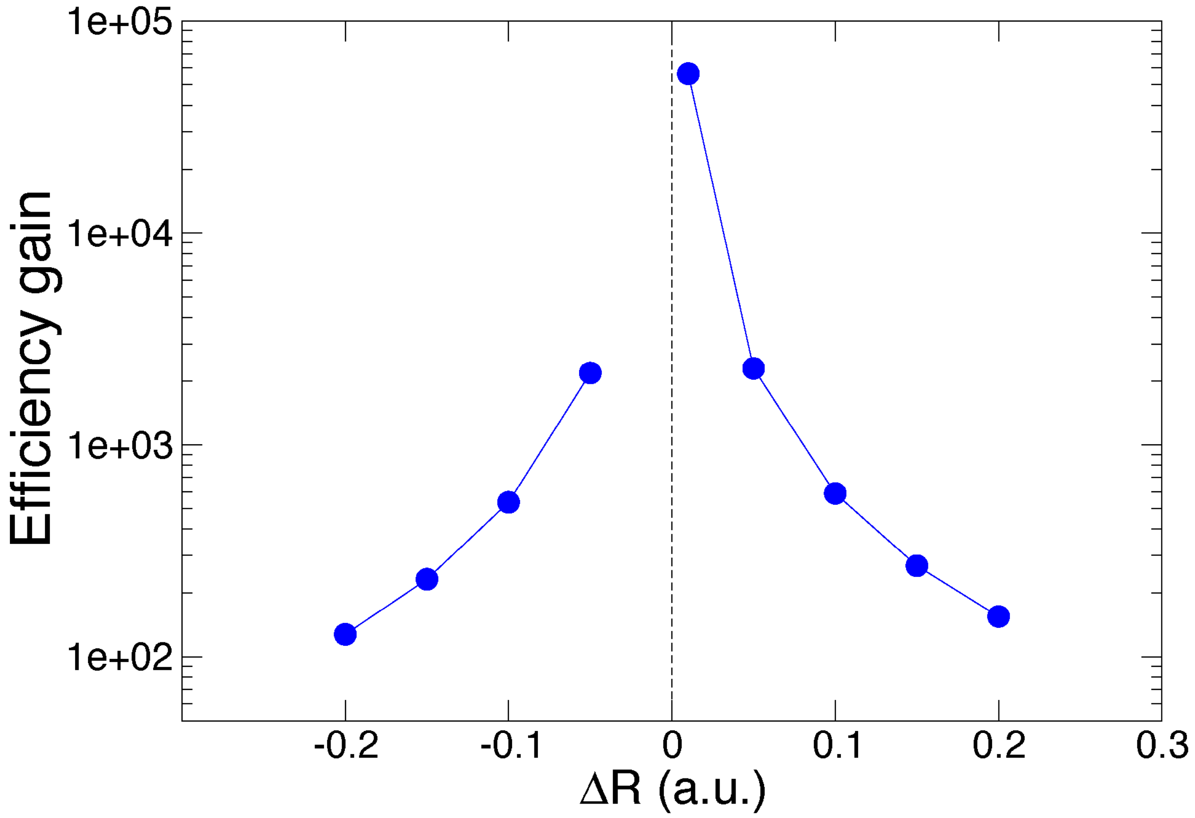
E_0 at expt. equilibrium bond length $R_{\text{exp}}^{\text{eq}} = 3.005$ a.u.

E at stretched bond length by $\Delta R = -0.2, \ldots, 0.2$

Compute $E - E_0$ from independent runs $\rightarrow \Delta E_{\text{ind}}$

from correlated sampling $\rightarrow \Delta E_{\text{corr}}$

$$\text{Efficiency gain} = \frac{\sigma^2(\Delta E_{\text{ind}})}{\sigma^2(\Delta E_{\text{corr}})}$$



Note: We used space-warp coordinate transformation

⋆ $\boxed{\text{This afternoon}}$

Compute bond length of B_2

1. Construct reference trial wave function

$$\Psi = \underbrace{\sum_i d_i\, D_i}_{\text{sum of determinants}} \times \mathcal{J} \quad \leftarrow \text{simple Jastrow factor}$$

Optimize Ψ by variance minimization

Choose experimental equilibrium bond length $R_{\text{exp}}^{\text{eq}}$

B_2: 1 det + simple Jastrow, $E_{\text{VMC}}^{\text{corr}}$ =28%

2. Secondary geometry wave functions

$\mathbf{R}_\alpha \quad \Psi$

$\mathbf{R}_\alpha^{\text{s}} \quad \Psi_{\text{s}} \quad \Rightarrow$ What do we use for Ψ_{s}?

Simple choice: Recenter the wave function at new nuclear positions and keep the same parameters

$$\Psi_{\text{s}}(\mathbf{R},\mathbf{R}_\alpha^{\text{s}}) = \Psi(\mathbf{R},\mathbf{R}_\alpha^{\text{s}},\mathbf{p})$$

Better choice: $\Psi_{\text{s}}(\mathbf{R},\mathbf{R}_\alpha^{\text{s}}) = \Psi(\mathbf{R},\mathbf{R}_\alpha^{\text{s}},\mathbf{p}_{\text{s}})$ with re-optimized parameters (smaller fluctuations in ΔE)

3. Space-warp coordinate transformation

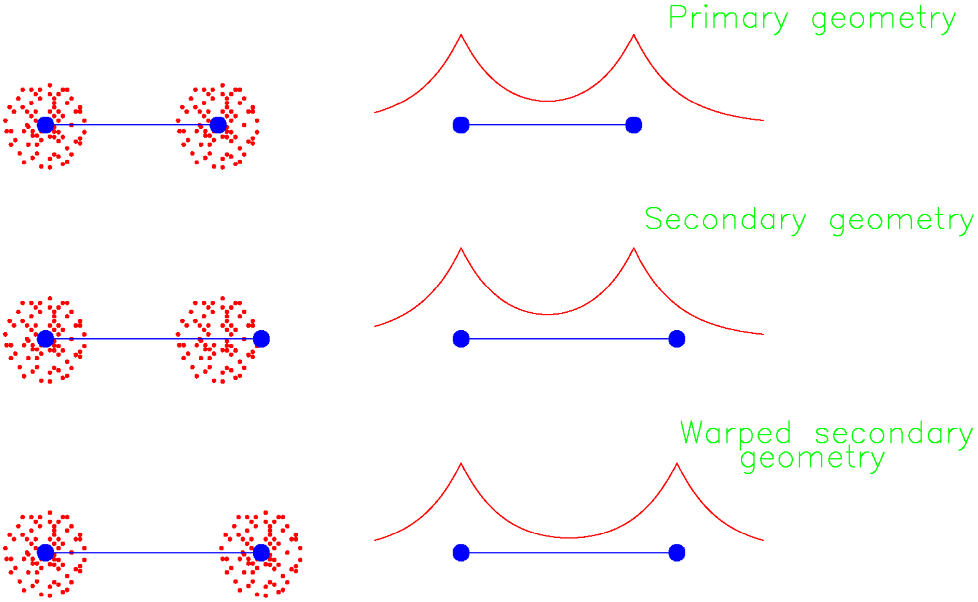
Primary geometry $\mathcal{H}$ Ψ $\mathbf{R}_\alpha$

Secondary geometry $\mathcal{H}_\mathrm{s}$ Ψ_s $\mathbf{R}^\mathrm{s}_\alpha$

↑

nuclear positions

We sample MC configurations from Ψ^2



Electrons close to a nucleus move
almost rigidly with the nucleus

Primary geometry $\mathcal{H}$ Ψ $\mathbf{R}_\alpha$ $\mathbf{R} = (\mathbf{r}_1, \dots, \mathbf{r}_N)$

Secondary geometry $\mathcal{H}_\mathrm{s}$ Ψ_s $\mathbf{R}^\mathrm{s}_\alpha$ $\mathbf{R}^s = (\mathbf{r}^s_1, \dots, \mathbf{r}^s_N)$

How do we map primary to secondary walker?

$$\begin{aligned}
\mathbf{r}_i^{\mathrm{s}} &= \mathbf{r}_i + \sum_{\alpha=1}^{N_{\mathrm{atom}}} (\mathbf{r}_\alpha^{\mathrm{s}} - \mathbf{r}_\alpha)\, \omega_\alpha(\mathbf{r}_i) \\
\omega_\alpha(\mathbf{r}_i) &= \frac{F(|\,\mathbf{r}_i - \mathbf{r}_\alpha|)}{\sum_{\beta=1}^{N_{\mathrm{atom}}} F(|\,\mathbf{r}_i - \mathbf{r}_\beta|)}
\end{aligned}$$

e.g. with $F(r) = r^{-\kappa}$ and $\kappa = 4$

Energy difference with space-warp transformation

$$E_{\mathrm{s}} - E = \frac{1}{N_{\mathrm{conf}}} \sum_{i=1}^{N_{\mathrm{conf}}} \left\{ \frac{\mathcal{H}_{\mathrm{s}} \Psi_{\mathrm{s}}(\mathbf{R}_i^{\mathrm{s}})}{\Psi_{\mathrm{s}}(\mathbf{R}_i^{\mathrm{s}})} w_i - \frac{\mathcal{H}\Psi(\mathbf{R}_i)}{\Psi(\mathbf{R}_i)} \right\}$$

↗ stretched coordinates

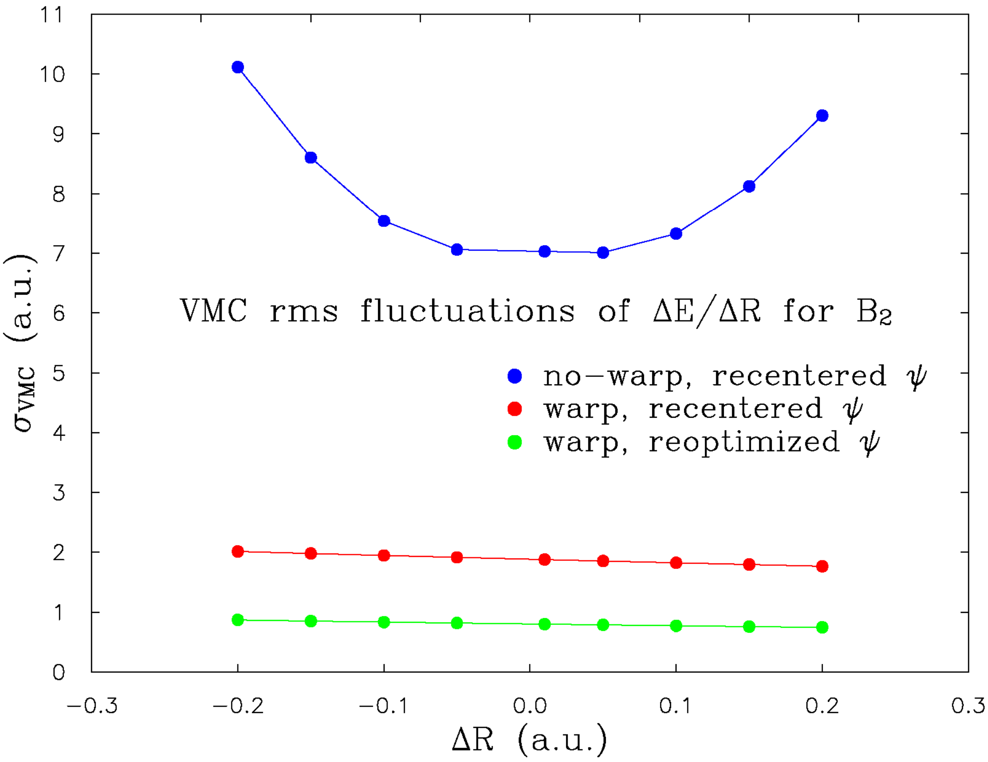
↑ primary coordinates

$$w_i = \frac{\left|\Psi_{\mathrm{s}}(\mathbf{R}_i^{\mathrm{s}})/\Psi(\mathbf{R}_i)\right|^2 \boxed{J(\mathbf{R}_\mathrm{i})}}{\frac{1}{N_{\mathrm{conf}}} \sum_{j=1}^{N_{\mathrm{conf}}} \left|\Psi_{\mathrm{s}}(\mathbf{R}_j^{\mathrm{s}})/\Psi(\mathbf{R}_j)\right|^2 \boxed{J(\mathbf{R}_\mathrm{j})}}$$

$J(\mathbf{R})$ Jacobian of transformation $\mathbf{R} \longrightarrow \mathbf{R}^{\mathrm{s}}$

Example: Bond length of B_2

- Multi-determinants + e-e-n Jastrow function
- $E^{\text{corr}}_{\text{VMC}} = 83\%$ at $\mathbf{R}^{\text{eq}}_{\text{exp}} = 3.005$ a.u.
- Root-mean-square fluctuations of $\dfrac{\Delta E}{\Delta R}$



- Error in bond length

(a.u.)	RHF	LDA	GGA	VMC	DMC
ΔR_e	0.086	0.025	0.042	0.018(2)	0.002(2)

4. Compute forces and bond length for B_2

Interatomic forces and geometry optimization

One possible route: correlated sampling

What about Hellman-Feynman theorem?

$\mathcal{H}(\lambda)$, λ parameter (nuclear coordinates)

$$E(\lambda) = \frac{\langle \Psi(\lambda) \left| \mathcal{H}(\lambda) \right| \Psi(\lambda) \rangle}{\langle \Psi(\lambda) | \Psi(\lambda) \rangle}$$

$$\frac{\mathrm{d}E(\lambda)}{\mathrm{d}\lambda} = \frac{\left\langle \Psi(\lambda) \left| \frac{\mathrm{d}\mathcal{H}(\lambda)}{\mathrm{d}\lambda} \right| \Psi(\lambda) \right\rangle}{\langle \Psi(\lambda) | \Psi(\lambda) \rangle}$$

True if a) $\Psi(\lambda)$ is an eigenstate

or b) $\Psi_\alpha(\lambda)$ minimizes the energy wrt α

Problems with Hellman-Feynman forces in QMC

- Ψ_T does not minimize the VMC energy:
  Hellman-Feynman $\Rightarrow$ systematic error in VMC
- Ψ_T does not minimize the DMC energy:
  Hellman-Feynman $\Rightarrow$ systematic error in DMC
- Large fluctuations: infinite for all electrons!
  $\longrightarrow$ Reduced variance method by Caffarel

Customary practice: use DFT or QC geometries

References

- Early applications of variational Monte Carlo:

  W.L. McMillan, Phys. Rev. **138**, A4422 (1965).

  D. Ceperley, G.V. Chester, M.H. Kalos, Phys. Rev. B **16**, 3081 (1977).

- VMC and choice of transition matrix:

  C.J. Umrigar, Phys. Rev. Lett. **71**, 408 (1993).

  C.J. Umrigar in "Quantum Monte Carlo methods in Physics and Chemistry", edited by M.P. Nightingale and C.J. Umrigar (Nato Science Series, 1999).

- Cusp conditions:

  R. T. Pack and W. Byers Brown, J. Chem. Phys. **45**, 556 (1966).

- Spin-assigned wave function, spin contamination:

  C-J. Huang, C. Filippi, C.J. Umrigar, J. Chem. Phys. **108**, 8838 (1998).

- Jastrow factor and role of e-e-e-n terms:

  C-J. Huang, C.J. Umrigar, M.P. Nightingale, J. Chem. Phys. **107**, 3007 (1997).

- Variance minimization:

  R.L. Coldwell, Int. J. Quantum Chem. Symp. **11**, 215 (1977).

  C.J. Umrigar, K.G. Wilson, J.W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).

  C. Filippi and C.J. Umrigar, J. Chem. Phys. **105**, 213 (1996).

- Fluctuation potential method:

  C. Filippi, S. Fahy, J. Chem. Phys. **112**, 3523 (2000).

  F. Schautz, S. Fahy, J. Chem. Phys. **116**, 3533 (2002).

  D. Prendergast, D. Bevan, S. Fahy, Phys. Rev. B **66**, 155104 (2002)

- Correlated sampling for interatomic forces:

  C. Filippi, C.J. Umrigar, Phys. Rev. B **61**, R16291 (2000).

- Hellman-Feynman forces in QMC:

  R. Assaraf, M. Caffarel, Phys. Rev. Lett. **83**, 4682 (1999); J. Chem. Phys. **113**, 4028 (2000).